

### **Paper 3**

**Title: Power Analysis for Cross Level Mediation in CRTs**

**Authors and Affiliations:**

Ben Kelcey

University of Cincinnati

**Background / Context:**

A common design in education research for interventions operating at a group or cluster level is a cluster randomized trial (CRT) (Bloom, 2005). In CRTs, intact clusters (e.g., schools) are assigned to treatment conditions rather than individuals (e.g., students) and are frequently an effective way to study interventions because they permit researchers to accommodate existing school structures and the multilevel nature of many educational interventions (Spybrook & Raudenbush, 2009).

A principal consideration in the design of CRTs is the power with which a study can detect effects if they exist (Hedges & Hedberg, 2007; Raudenbush, 1997). To assist in the design of CRTs, literature has developed statistical power formulas for a range of cluster based designs (e.g., Schochet, 2011; Hedges, Hedberg, & Kuyper, 2012). Based on these frameworks, literature has also compiled empirical estimates of the design parameters (e.g., intraclass correlations) to describe reasonable values for the design of studies (e.g., Jacobs, Zhu & Bloom, 2010).

Although CRTs provide rigorous evidence concerning the effect of the treatment on the outcome, there is a growing recognition of the need to also investigate the mechanisms through which the treatment is presumed to work (e.g., Hong, 2012). One common approach is to establish a sequence of structural relationships through a mediation analysis. Mediation analyses examine the extent to which a treatment has an indirect effect on an outcome by examining changes in a mediator produced by exposure to a treatment and, in turn, how these changes in the mediator are associated with changes in outcomes. Despite the importance of mediation analyses and the value of well-planned studies, literature is incomplete in describing both power formulas for cross-level mediation effects in CRTs and reasonable values for the formula parameters.

**Purpose / Objective / Research Question / Focus of Study:**

In this study, we developed closed form analytic expressions for the test statistics of the mediation effect in three level CRT designs examining cross-level (e.g., 3-2-1) mediation. Using this test statistic, we estimate the power to detect cross-level mediation effects under different scenarios to assist researchers in the planning and design of multilevel mediation studies.

To fix ideas, consider an example where the intervention is a professional development program to increase pedagogical knowledge and it is assigned at random to intact schools (e.g., Garet et al., 2008; 2011). In theory, exposure to the professional development program (treatment) should increase teachers' pedagogical content knowledge (mediator) which in turn increases students' achievement levels (outcome). We might describe the study as a 3-2-1 cross-level mediation design because its hypothesized that a treatment assigned at the school level (level 3) impacts a classroom level (level 2) mediator which in turn impacts a student level (level 1) outcome. The focus of this study was to develop a framework to assess the power to detect cross-level mediation for this particular three level CRT design.

**Significance / Novelty of study:**

In detecting cross-level (3-2-1) mediation effects in three level CRTs, literature has largely relied on simulation or only the power to detect effects along single paths (e.g.,  $a$  or  $b$  paths) (e.g., Schochet, 2011). By developing closed form expressions of the test statistic for the product of cross-level effects, we provide tools for researchers to carefully plan (e.g., the power of) their studies so that they are well-positioned to detect mediation effects if they exist.

**Statistical, Measurement, or Econometric Model:**

We investigated two situations in our study. In the first situation, we examined 3-2-1 cross-level mediation effects as they pertain to the second (teacher) level only. In contrast, in the second situation we examined estimation of both level two (teacher) and level three (school)

mediation effects. We assumed a fully balanced design (i.e., equal number of students/classroom, classrooms/school, and equal allocation of school treatment and control conditions) and did not consider variables that might confound the mediator-outcome relationship. Although the treatment-mediator path is unconfounded because of random assignment, the mediator-outcome path may be confounded because levels of the mediator were not randomly assigned. We briefly discuss this and other limitations and extensions in the Conclusion section.

*Situation 1.* Mediation effects have been conventionally formulated using a linear structural equation framework (e.g., Baron & Kenny, 1986) (Figure 1). For instance, consider the following set of linear equations

$$\begin{aligned} M_{jk} &= \beta_{0j}^M + u_{jk}^M \\ \beta_{0j}^M &= \gamma_{00}^M + aT_k + r_{0k}^M \end{aligned} \quad (1) \quad \text{and} \quad \begin{aligned} Y_{ijk} &= \pi_{0jk}^Y + \varepsilon_{ijk}^Y \\ \pi_{0jk}^Y &= \beta_{00k}^Y + bM_{jk} + u_{0jk}^Y \\ \beta_{00k}^Y &= \gamma_{000}^Y + \gamma_{001}^Y T_k + r_{00k}^Y \end{aligned} \quad (2)$$

In expression (1), the first equation describes the relationships of teachers within schools where  $M_{jk}$  is a continuously distributed mediator for teacher  $j$  in school  $k$ ,  $\beta_{0j}^M$  is the intercept for the mediator equation and  $u_{jk}^M$  is the normally distributed error term with variance  $\hat{\sigma}_{u^M}^2$ . The second equation describes school-level relationships where  $\gamma_{00}^M$  is the grand intercept,  $T_k$  is the treatment with associated coefficient  $a$  describing its relationship with the mediator, and  $r_{0k}^M$  is a normally distributed error term with variance  $\hat{\sigma}_{r_0^M}^2$ . In expression (2),  $Y_{ijk}$  is the outcome for student  $i$  served by teacher  $j$  in school  $k$ ,  $\pi_{0jk}^Y$  is the intercept,  $\varepsilon_{ijk}^Y$  is the normally distributed error term with variance  $\hat{\sigma}_{\varepsilon^Y}^2$ ,  $\beta_{00k}^Y$  is the teacher intercept,  $M_{jk}$  is the mediator with associated coefficient  $b$  capturing the conditional association between the mediator and the outcome (which has coefficient  $\gamma$ ),  $u_{0jk}^Y$  is a normally distributed teacher random effect term with variance  $\hat{\sigma}_{u_0^Y}^2$ ,  $\gamma_{000}^Y$  is the school level intercept,  $T_k$  is the treatment indicator with associated coefficient  $\gamma_{001}^Y$ , and  $r_{00k}^Y$  is a normally distributed school random effect term with variance  $\hat{\sigma}_{r_0^Y}^2$ .

One approach to estimating the indirect effect using this system of equations involves fitting a sequence of multilevel models to expressions (1) and (2) (e.g., Zhang, Zyphur, & Preacher, 2009) so that the mediation effect is the product of the direct effect of the treatment of interest on the mediator ( $a$  in expression (1)) by the direct effect of the mediator on the outcome controlling for the treatment ( $b$  in expression (2)).

The asymptotic variance of the (single) mediation effect can be estimated using the Delta method (Sobel, 1982; Bollen, 1987; 1989),

$$\sigma_{ab}^2 = b^2 S_a^2 + a^2 S_b^2 \quad (3)$$

In turn, the significance of the mediation effect in large samples can be approximated using

$$z_m = \frac{ab}{\sqrt{b^2 \sigma_a^2 + a^2 \sigma_b^2}} \quad (4)$$

where  $z_m$  has an approximate normal distribution in large samples (Sobel, 1982). In the context of a three-level CRT with treatment assigned to schools and a level two mediator, the variances of the  $a$  and  $b$  estimated paths are

$$\hat{\sigma}_a^2 = \frac{\hat{\sigma}_{u^M}^2 + n_2 \hat{\sigma}_{r_0^M}^2}{n_2 n_3 \hat{\sigma}_T^2} \quad \text{and} \quad \hat{\sigma}_b^2 = \frac{\hat{\sigma}_{\varepsilon^Y}^2 + n_1 \hat{\sigma}_{u_0^Y}^2}{n_1 n_2 n_3 \hat{\sigma}_M^2} \quad (5)$$

where  $n_1$ ,  $n_2$ , and  $n_3$  are the student, teacher, and school sample sizes. We can write the test statistic for the mediation effect as

$$z_m = \frac{ab}{\sqrt{b^2 \frac{\hat{\sigma}_{u^M}^2 + n_2 \hat{\sigma}_{r_0^M}^2}{n_2 n_3 \hat{\sigma}_T^2} + a^2 \frac{\hat{\sigma}_{\varepsilon^Y}^2 + n_1 \hat{\sigma}_{u_0^Y}^2}{n_1 n_2 n_3 \hat{\sigma}_M^2}}} \quad (6)$$

As a result, similar to power analyses for detecting main effects in CRTs, given a priori estimates of the standardized variance components and the expected magnitude of the path coefficients, we can estimate the power of a given design.

*Situation 2.* In the second situation we examined scenarios in which the within-group effects differ from the between-group effects (Zhang, Zyphur, & Preacher, 2009). We adapt equations (1) and (2) to now include cluster (school) level mediator means such that

$$\begin{aligned} M_{jk} &= \beta_{0j}^M + u_{jk}^M & Y_{ijk} &= \pi_{0jk}^Y + \varepsilon_{ijk}^Y \\ \beta_{0j}^M &= \gamma_{00}^M + aT_k + r_{0k}^M & \pi_{0jk}^Y &= \beta_{00k}^Y + b_W(M_{jk} - \bar{M}_{\bullet k}) + u_{0jk}^Y \\ & & \beta_{00k}^Y &= \gamma_{000}^Y + \gamma_{001}^Y T_k + b_B \bar{M}_{\bullet k} + r_{00k}^Y \end{aligned} \quad (7) \quad \text{and} \quad (8)$$

Here we continue with the aforementioned notation and introduce  $b_W$  as the within-group effect,  $b_B$  as the between-group effect, and  $\bar{M}_{\bullet k}$  as mean of the mediator for school  $k$  such that mediator values for teachers are centered on their school means (i.e., group-mean centered). Because group-mean centered variables at the teacher level are independent of the school level means, the power for the within-group (teacher) mediation effect ( $b_W$ ) can be estimated using the steps outlined above.

To estimate power for the between-group (school) mediation effect ( $b_B$ ), we begin by noting that an unbiased estimate of  $a$  in equation 8 can had using an OLS estimator on the aggregated data so that

$$\bar{M}_k = \gamma_{00}^M + aT_k + r_{0k}^{M*} \quad (9)$$

An unbiased estimate of  $a$  and its standard error can be obtained using

$$\hat{a} = \frac{\sigma_{\bar{M}_k}}{\sigma_{T_k}} (\rho_{\bar{M}_k T_k}) \quad \sigma_{\hat{a}} = \frac{\sigma_{\bar{M}_k}}{\sigma_{T_k}} \times \frac{\sqrt{1 - r_{\bar{M}_k T_k}^2}}{\sqrt{n_3 - 2}} \quad t_{\hat{a}} = \frac{\rho_{\bar{M}_k T_k}}{\sqrt{1 - \rho_{\bar{M}_k T_k}^2}} \quad (10)$$

where  $\sigma_{\bar{M}_k}$  and  $\sigma_{T_k}$  are the standard deviations of the aggregated mediator and treatment variables,  $\rho_{\bar{M}_k T_k}$  is their correlation. Similarly, we can estimate  $b_B$  using (note that although  $M_{jk}$  and  $\bar{M}_{\bullet k}$  are independent,  $T_k^Y$  and  $\bar{M}_{\bullet k}$  are not independent)

$$\bar{Y}_k = \gamma_{000}^Y + \gamma_{001}^Y T_k^Y + b_B \bar{M}_{\bullet k} + r_{00k}^{Y*} \quad (11)$$

Where  $b_B$  and its standard error are

$$\hat{b}_B = \left( \frac{\sigma_{\bar{Y}_k}}{\sigma_{\bar{M}_k}} \right) \frac{\rho_{\bar{Y}_k \bar{M}_k} - \rho_{\bar{Y}_k T_k} \rho_{\bar{M}_k T_k}}{1 - \rho_{\bar{M}_k T_k}^2}, \quad \sigma_{\hat{b}_B} = \left( \frac{\sigma_{\bar{Y}_k}}{\sigma_{\bar{M}_k}} \right) \sqrt{\left( \frac{1 - \rho_{\bar{Y}_k \bar{M}_k}^2 - 2\rho_{\bar{Y}_k \bar{M}_k} \rho_{\bar{Y}_k T_k} \rho_{\bar{M}_k T_k}}{1 - \rho_{\bar{M}_k T_k}^2} \right) \left( \frac{1}{n_3 - 3} \right) \left( \frac{1}{1 - \rho_{\bar{M}_k T_k}^2} \right)} \quad \text{and} \quad t_{\hat{b}_B} = \frac{\rho_{\bar{Y}_k \bar{M}_k}}{\sqrt{1 - \rho_{\bar{Y}_k \bar{M}_k}^2}} \quad (12)$$

where  $\rho_{\bar{Y}_k \bar{M}_k | T_k}$  is the partial correlation such that

$$\rho_{\bar{Y}_k \bar{M}_k | T_k} = \frac{\rho_{\bar{Y}_k \bar{M}_k} - \rho_{\bar{Y}_k T_k} \rho_{\bar{M}_k T_k}}{\sqrt{(1 - \rho_{\bar{Y}_k T_k}^2)(1 - \rho_{\bar{M}_k T_k}^2)}} \quad (13)$$

Next we note based on above the statistical significance of the between mediation effect,  $ab_B$ , denoted as  $z_{ab_B}$ , can be estimated using

$$z_{ab_B} = \frac{ab_B}{\sqrt{b_B^2 \sigma_a^2 + a^2 \sigma_{b_B}^2}} \Rightarrow z_{ab_B}^{-2} = \frac{\sigma_{b_B}^2 a^2}{a^2 b_B^2} + \frac{\sigma_a^2 b_B^2}{a^2 b_B^2} \quad (14)$$

Similarly, we can write the  $t$ -values of the  $a$  and  $b_B$  parameters in (10) and (12) as

$$t_a = \frac{a}{\sigma_a} \Rightarrow t_a^{-2} = \frac{\sigma_a^2 b_B^2}{a^2 b_B^2} \quad \text{and} \quad t_{b_B} = \frac{b_B}{\sigma_{b_B}} \Rightarrow t_{b_B}^{-2} = \frac{\sigma_{b_B}^2 a^2}{a^2 b_B^2} \quad (15)$$

With some algebra we find that (assuming positive mediation)

$$z_{ab_B} = \frac{t_a t_{b_B}}{\sqrt{t_a^2 + t_{b_B}^2}} \quad (16)$$

Using expressions (10) and (12) we can rewrite this as

$$z_{ab_B} = \frac{\left( \frac{\rho_{\bar{M}_k T_k}}{\sqrt{1 - \rho_{\bar{M}_k T_k}^2}} \right) \left( \frac{\rho_{\bar{Y}_k \bar{M}_k T_k}}{\sqrt{1 - \rho_{\bar{Y}_k \bar{M}_k T_k}^2}} \right)}{\sqrt{\frac{\rho_{\bar{M}_k T_k}^2}{1 - \rho_{\bar{M}_k T_k}^2} + \frac{\rho_{\bar{Y}_k \bar{M}_k T_k}^2}{1 - \rho_{\bar{Y}_k \bar{M}_k T_k}^2}}} \quad (17)$$

This expression can be used to estimate the power with which we can detect a between group mediation effect if it exists. We could also parameterize the expressions in terms of  $a$  and  $b_B$  using, for example the following relationships

$$\hat{b}_B = \left( \frac{\sigma_{\bar{Y}_k}}{\sigma_{\bar{M}_k}} \right) \frac{\rho_{\bar{Y}_k \bar{M}_k} - \rho_{\bar{Y}_k T_k} \rho_{\bar{M}_k T_k}}{1 - \rho_{\bar{M}_k T_k}^2} \quad \text{and} \quad \rho_{\bar{Y}_k \bar{M}_k T_k} = \frac{\rho_{\bar{Y}_k \bar{M}_k} - \rho_{\bar{Y}_k T_k} \rho_{\bar{M}_k T_k}}{\sqrt{(1 - \rho_{\bar{Y}_k T_k}^2)(1 - \rho_{\bar{M}_k T_k}^2)}} \quad (18)$$

### Usefulness / Applicability of Method:

Using data from a large study of many different professional development programs, we applied the method to design studies examining the extent to which teachers' knowledge mediates the relationship between participation in a professional development program and students' cognitive development. Our application examines designs that achieve 80% power based on design parameters (e.g., ICC, effect size) from this study as well as other studies.

### Findings / Results:

Our results (not presented due to space constraints) suggested that studies investigating cross-level mediation using CRTs will typically need large sample sizes or large mediation effects to have an 80% chance of detecting mediation effects if they exist.

### Conclusions:

The method developed allows researchers to plan and design more effective studies of cross-level mediation. Our analysis of 3-2-1 mediation involving professional development, teacher knowledge, and student achievement underscored the importance of carefully planning and designing studies to detect effects. Studies with typical CRT sample sizes (e.g., 40 schools) will generally not be sufficient to detect mediation in this context. The framework we developed has several limitations. For instance, our framework assumes linearity, large sample approximations, sequential ignorability, that mediation is not moderated, mediator slopes do not randomly vary across schools, and it does not include latent variables. Subsequent works addresses some of these shortcomings.

## Appendices

### Appendix A. References

- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Bloom, H. (2005). Randomizing groups to evaluate place-based programs. In H.S. Bloom (Ed.), *Learning more from social experiments: Evolving analytic approaches* (pp. 115–172). New York: Russell Sage Foundation
- Bollen, K. A. (1987). Total, direct, and indirect effects in structural equation models. *Sociological Methodology*, 17, 37-69.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Garet, M.S, Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., Uekawa, K., Falk, A., Bloom, H., Doolittle, F., Zhu, P., Szejnbert & L. Silverberg, M. (2008). The Impact of Two professional development Interventions on Early Reading Instruction and Achievement. U.S. Department of Education Report NCEE 2008-4030.
- Garet, M.S., Wayne, A.J., Stancavage, F., Taylor, J., Eaton, M., Walters, K., Song, M., Brown, S., Hurlburt, S., Zhu, P., Sepanik, S & Doolittle, F. (2011). Middle School Mathematics professional development Impact Study: Findings After the Second Year of Implementation. U.S. Department of Education Report NCEE 2011-4024.
- Hedges, L., Hedberg, E., & Kuyper, A., (2012). The variance of intraclass correlations in three- and four-level models. *Educational and Psychological Measurement*, 1-17.
- Hedges, L., & Hedberg, E. (2007). Intraclass Correlation Values for Planning Group-Randomized Trials in Education. *Educational Evaluation and Policy Analysis*, 29,1, pp. 60-87.
- Hong, G. (2012). Editorial Comments. *Journal of Research on Educational Effectiveness*, 5, 213-214.
- Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods*, 2(2), 173.
- Schochet, P. (2011). Do typical RCTs of education interventions have sufficient statistical power for linking impacts on teacher practice and student achievement outcomes? *Journal of Educational and Behavioral Statistics*, 36, 4, 441-471.
- Sobel, M. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology* 13 290–321.
- Spybrook, J. & Raudenbush, S. W. (2009). An examination of the precision and technical accuracy of the first wave of group randomized trials funded by the Institute of Education Sciences. *Educational Evaluation and Policy Analysis*. Vol. 31, No. 3, pp 298 – 318.

Zhang, Z., Zyphur, M., & Preacher, K. (2009). Testing multilevel mediation using hierarchical linear models. *Organizational research Methods*, 12, 4, 695-719.

## Appendix B. Tables and Figures

Figure 1

